

Analisis Pola Hubungan Antara Konsumsi Listrik Dengan Temperatur dan Fitur Geografi Menggunakan *Association Rule Mining*

Haris Prasetyo, Imam Mukhlash, dan Nurul Hidayat

Departemen Matematika, Fakultas Matematika Komputasi dan Sains Data, Institut Teknologi Sepuluh Nopember (ITS)

email: imamm@matematika.its.ac.id

Abstrak—Perkembangan zaman semakin mempengaruhi pola hidup masyarakat yang mana menyebabkan penggunaan listrik mengalami peningkatan. Sementara itu biaya pengadaan tenaga listrik serta distribusinya juga membutuhkan biaya yang tidak kecil. Maka dari itu diperlukan kajian lebih lanjut tentang pengetahuan profil-profil penggunaan listrik tiap wilayah agar pendistribusiannya efektif. Untuk menggali pengetahuan ini diperlukan teknik *data mining*. Dalam penelitian ini digunakan algoritma *fp-growth* yaitu dengan membuat *fp tree*, *conditional pattern base* lalu menemukan *pattern* atau pola yang terbentuk dan diubah menjadi suatu *rule*. Dari data selama satu tahun (September 2015-Agustus 2016) daerah Sidoarjo, didapat bahwa wilayah yang jauh dengan area perikanan cenderung tinggi dalam penggunaan listriknya. Pemakaian listrik juga tinggi untuk daerah yang dekat dengan pertanian dan jalan besar. Lalu ketika temperatur berkisar 28,5°-29,5° pemakaian listrik tergolong rendah

Kata Kunci—Konsumsi Listrik, *Data Mining*, Asosiasi, *FP-Growth*.

I. PENDAHULUAN

KEBERLANJUTAN energi selalu menjadi topik perbincangan yang menarik untuk terus dikaji. Salah satunya yaitu energi listrik. Saat ini banyak tersebar slogan-slogan seperti ‘*save energy*’ dan lain sebagainya. Hal ini menandakan bahwa kesadaran masyarakat akan keberlanjutan energi semakin meningkat. Hal ini dipengaruhi oleh perkembangan zaman yang semakin maju yang menyebabkan teknologi-teknologi yang digunakan manusia juga semakin canggih. Namun sebagian besar teknologi tersebut membutuhkan energi listrik untuk penggunaannya yang akhirnya akan berdampak pada peningkatan penggunaan energi listrik. Sementara itu sebagian besar bauran energi mix untuk pengadaan tenaga listrik bergantung pada sumber daya yang relatif lama atau tidak bisa diperbarui. Di sisi lain biaya pengadaan serta distribusinya juga membutuhkan biaya yang tidak kecil. Maka dari itu diperlukan kajian lebih lanjut tentang pengetahuan profil-profil penggunaan listrik tiap wilayah agar pendistribusiannya efektif. Untuk menggali pengetahuan ini diperlukan teknik data mining. Pengetahuan yang diperoleh

berupa suatu *rule* yang berkaitan dengan penggunaan listrik. Informasi seperti itu bisa menjadi bahan pertimbangan dalam menentukan distribusi energi listrik agar lebih efektif.

Dalam penelitian ini, algoritma yang digunakan yaitu *fp-growth*, salah satu algoritma asosiasi yang digunakan untuk menemukan pola hubungan antar dua hal atau lebih. Algoritma ini merupakan pengembangan dari metode asosiasi yang sebelumnya yang dikenal dengan algoritma *apriori*. Berbeda dengan algoritma *apriori* yang butuh berkali-kali untuk memindai data dari basis data, *fp-growth* butuh dua kali pemindaian data. Langkah selanjutnya yaitu membentuk *fp tree* sebagai pengganti basis data yang akan dipindai untuk yang kesekian kalinya jika pada algoritma *apriori*. Langkah selanjutnya dibentuklah *conditional pattern base*, lalu menemukan *frequent pattern* atau pola yang sering muncul. Dan diakhiri dengan *generate rule*. *Rule* ini merupakan kombinasi dari beberapa item yang sesuai dengan *frequent pattern* tersebut.

Data yang digunakan adalah data salah satu kota di Jawa timur yaitu kabupaten Sidoarjo. Data tersebut antara lain temperatur rata-rata per hari, data jarak setiap wilayah dengan fitur geografi (kawasan pertanian, perikanan, dan jalan besar), serta data penggunaan listrik per bulan dari masing-masing wilayah tersebut. Data temperatur nantinya akan diubah menjadi data bulanan dengan menggunakan salah satu rumus agregasi yaitu *average*. Data jarak setiap wilayah dengan fitur geografi didapat dari pengukuran peta menggunakan software GIS.

II. TINJAUAN PUSTAKA

A. Penelitian Terkait

Penulis memulai penelitian ini dengan terlebih dahulu melakukan studi kepustakaan dari penelitian-penelitian dan sumber-sumber lain. Penelitian tersebut membahas tentang topik yang terkait dengan penelitian penulis, antara lain mengenai masalah yang diangkat serta mengenai metode yang digunakan penulis.

Dengan menggunakan sebuah formula yang disebut *ideal curve* Jovanovic, S dkk [1] melakukan sebuah penelitian tentang akibat atau dampak dari suatu perubahan temperatur udara per hari terhadap konsumsi listrik pada kota Kragujevac (Serbia). Dalam penelitian tersebut dinyatakan bahwa

temperatur rata-rata udara tiap hari merupakan sebuah parameter iklim yang paling berpengaruh dan memberikan dampak yang luas. Peningkatan dan penurunan konsumsinya tergantung pada simpangan dari temperatur rata-rata per hari. Pada musim dingin terjadi peningkatan konsumsi listrik sebagai hasil dari proses pemanasan baik di daerah pemukiman maupun perkantoran. Selama musim panas yang panjang dan ekstrem juga terdapat suatu kecenderungan peningkatan konsumsi listrik karena pendingin udara.

Telah dilakukan sebuah penelitian konsumsi listrik per jam terhadap pemukiman Norwegia [2]. Pada penelitian tersebut dilakukan analisa untuk mengetahui seberapa besar sistem pemanasan mempengaruhi konsumsi listrik per jam oleh setiap rumah. Data meteran listrik per jam, data cuaca, dan data respon dari survei rumah tangga digunakan, serta model multi regresi diaplikasikan untuk mengetahui dampak dari sistem pemanas ini.

Penelitian lain dilakukan di Indonesia untuk mengetahui karakteristik konsumsi listrik rumah tangga. Faktor yang menggiring konsumsi listrik tersebut ditentukan melalui survei lapangan dan diklasifikasikan menurut bagaimana penggunaannya. Hubungan antara konsumsi listrik dengan faktor yang mempengaruhi tersebut dievaluasi menggunakan analisa data multivariat. Penelitian tersebut dilakukan di dua kota berbeda untuk memperoleh perbandingan. Dan dari hasil yang diperoleh, disimpulkan bahwa kebijakan konservasi energi tidak bisa diperlakukan secara general melainkan harus lebih spesifik berdasarkan karakteristik yang diadopsi oleh masing-masing daerah [3].

Rathod dan Garg melakukan sebuah penelitian tentang hubungan antara konsumsi listrik dengan temperatur dan fitur geografi [4]. Penelitian ini mengambil data dari kota Shangli, India. Dalam penelitian ini mereka menggunakan algoritma apriori dan fitur geografi yang dipakai adalah sungai, sawah, lahan terbuka, dan jalan tol. Pada penelitian [1] dan [2] hanya dilakukan penelitian terkait temperatur. Sedangkan di penelitian [4], cakupannya sudah lebih luas yaitu menambah keterkaitannya dengan fitur geografi. Namun, algoritma yang digunakan yaitu apriori masih memiliki banyak kekurangan dan keterbatasan.

B. Data Mining

Aplikasi basis data sudah banyak diterapkan di berbagai bidang seperti manajemen industri, ilmu pengetahuan, administrasi pemerintah, dan masih banyak bidang yang lainnya. Hal ini menyebabkan data yang dihasilkan dari bidang-bidang tersebut sangat besar dan berkembang pesat. Dari keadaan seperti itu, muncul kebutuhan yaitu teknik seperti apa yang bisa digunakan untuk mengolah data tersebut sehingga bisa diperoleh informasi implisit yang berharga dari data-data tersebut. *Data Mining* adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam data. *Data Mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, *machine learning* untuk mengekstraksi dan informasi yang

bermanfaat dan pengetahuan yang terkait dari berbagai data besar [5].

C. Association Rule Mining

Association rule mining adalah pencarian aturan-aturan hubungan antar item dari suatu basis data transaksi atau basis data relasional. Dengan kata lain menemukan *frequent patterns*, asosiasi, korelasi, atau *causal structures* diantara himpunan item atau obyek dalam database transaksional, database relasional, dan penyimpanan lainnya.

Misal I adalah suatu himpunan set, maka suatu *rule* didefinisikan dengan $A \Rightarrow B$ dimana $A \subset I$, $B \subset I$, $A \cap B = \emptyset$. Dua parameter yang digunakan untuk mendapatkan *rule* yaitu:

1. *Support* (s), menyatakan probabilitas item-itemnya muncul dalam basis data atau bisa dikatakan suatu transaksi memuat $\{A \cup B\}$. $S = support(\{A \cup B\})$
2. *Confidence* (c), menyatakan probabilitas bahwa jika item-item pada ruas kiri muncul maka item-item pada ruas kanan juga akan muncul atau bisa dikatakan probabilitas bersyarat. $C = support(\{A \cup B\}) / support(\{A\})$.
3. *Lift*, nilai *lift* menunjukkan hubungan antar item. *Lift* bernilai <1 menunjukkan bahwa item-item tersebut berkorelasi negatif, dengan kata lain kemunculan salah satu item mempengaruhi hal yang sebaliknya pada kemunculan item lainnya. Untuk nilai $=1$ menandakan bahwa hubungan item-item adalah *independent*. Dan untuk nilai >1 menunjukkan bahwa item-item tersebut berkorelasi positif. Nilai *lift* bisa dicari dengan cara membagi *confidence* dengan *expected confidence*. Sedangkan *expected confidence* didapat dari $support(\{Y\}) / \text{jumlah record}$.

$$lift = \frac{\text{confidence}}{\text{expected confidence}} \quad (1)$$

$$conf = \frac{\text{support } A \cup B}{\text{support } A} \quad (2)$$

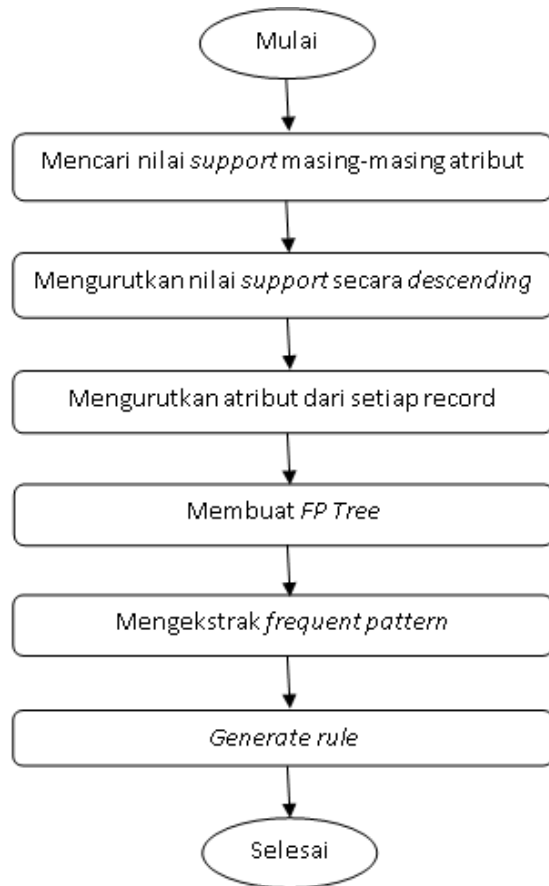
$$\text{expt conf} = \frac{\text{support } B}{\text{jumlah record}} \quad (3)$$

Salah satu algoritma dari *association rule* adalah *FP-Growth*. *FP-Growth* adalah perbaikan dari algoritma apriori, dibentuk untuk mengatasi kesulitan-kesulitan saat menggunakan apriori. *FP-Growth* menyederhanakan masalah-masalah yang ada ke dalam bentuk *FP-Tree*. Dalam *FP-Tree* ini, setiap *node* merepresentasikan sebuah item serta jumlahnya, dan setiap cabang merepresentasikan asosiasi yang berbeda-beda.

Dengan input berupa suatu *record* data dari basis data yang mengandung satu atau beberapa item. Dan outputnya adalah *frequent pattern*, yaitu suatu himpunan yang memuat satu atau lebih item dengan nilai *support* yang memenuhi besar *minimum support* yang ditentukan di awal.

Algoritma *FP-Growth*:

1. Meminindai data dan mencari nilai *support* dari masing-masing item. Dari semua data yang ada dalam basis data



Gambar 1. Alur berjalannya system.

akan dilakukan pemindaian untuk mengetahui frekuensi kemunculan dari masing-masing item tersebut.

- Urutkan masing-masing *support* dari terbanyak ke terendah. Setelah mengetahui frekuensi kemunculan atau nilai *support*, maka akan diurutkan dari yang tertinggi ke terendah.
- Membuat *FP-Tree* dari masing-masing transaksi. Akan dibentuk sebuah *tree* dengan cara menjadikan item dengan nilai *support* tertinggi sebagai akar (*root*) dan item dengan nilai *support* terendah sebagai daun (*leaf*).
- Ekstrak *frequent itemset*, dari bawah ke atas (dari daun ke akar). Dari *tree* yang sudah terbentuk akan dilakukan ekstrak *frequent itemset* dimulai dari daun ke akar.

Keuntungan *FP-Growth* diantaranya bisa meminimalisir pemindaian basis data, menyingkat dataset, tidak perlu *candidate generation*, dan lebih cepat dibandingkan dengan *apriori*.

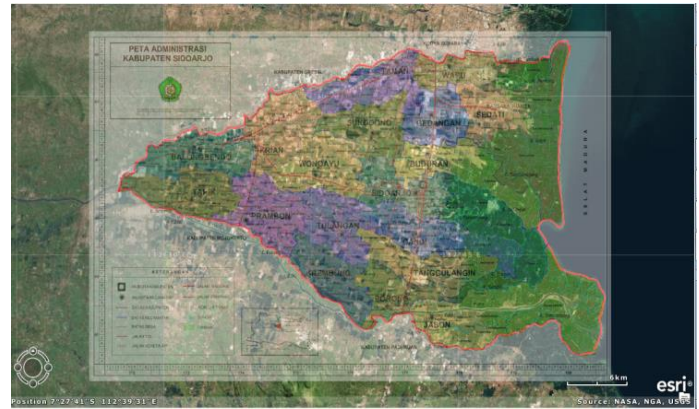
III. METODOLOGI PENELITIAN

A. Implementasi Data

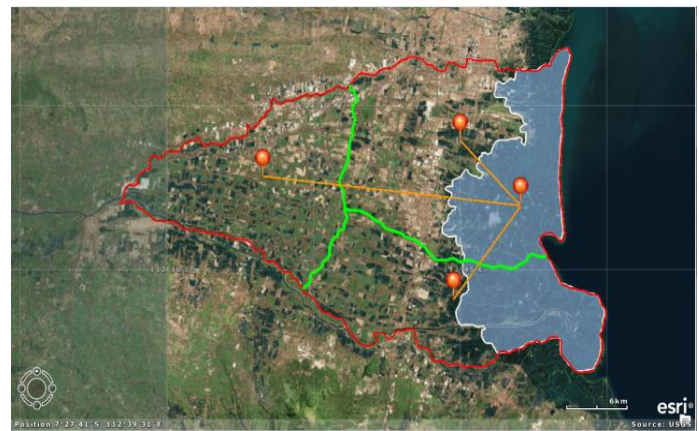
Data yang akan digunakan dalam sistem ini ada 3 macam, yaitu: data penggunaan listrik, data temperatur, dan data jarak antara fitur geografi dengan masing-masing wilayah.

1. Data penggunaan listrik

Data penggunaan listrik yang diperoleh dari data Kabupaten Sidoarjo meliputi jumlah pemakaian (kwh) dan jumlah pelanggan yang terbagi ke dalam 3 rayon, yaitu : sidoarjo kota,



Gambar 2. Membatasi wilayah sidoarjo sesuai peta administrasi yang tersedia.



Gambar 3. Mengukur jarak masing-masing rayon ke daerah perikanan.

krian, dan porong. Sebelum data ini diolah ke dalam sistem, perlu dilakukan proses untuk memperoleh bentuk data yang sesuai dengan rancangan sistem yang akan dibuat.

Pengolahan awal data untuk mendapatkan bentuk yang sesuai dengan kebutuhan, akan dikenakan fungsi *average* atau rata-rata untuk pemakaian masing-masing pelanggan per bulannya di setiap rayon.

$$Ave(\mu) = \frac{\text{pemakaian}}{\text{jumlah pelanggan}} \quad (4)$$

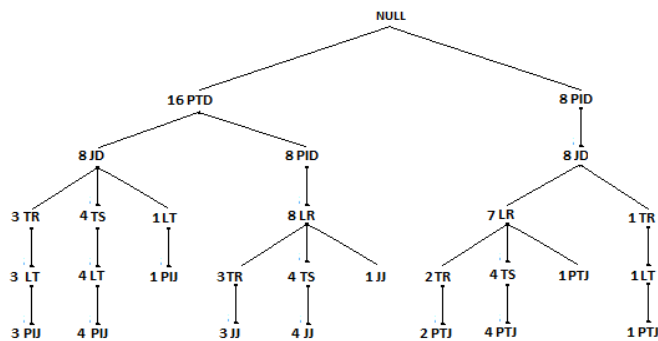
Setelah diperoleh pemakaian rata-rata untuk masing-masing rayon selama 12 bulan maka akan dibagi menjadi 2 kategori, yaitu rendah ($\leq \text{med}$) dan tinggi ($> \text{med}$).

$$\text{med} = \frac{\min(\mu_1, \mu_2, \mu_3, \dots) + \max(\mu_1, \mu_2, \mu_3, \dots)}{2} \quad (5)$$

2. Data temperatur

Data temperatur yang diperoleh untuk Kabupaten Sidoarjo meliputi nama stasiun, tanggal, dan suhu rata-rata. Karena stasiun yang tersedia hanya satu maka atribut tersebut diberlakukan untuk seluruh rayon.

Merujuk pada tujuan penelitian ini, yaitu mengetahui pola hubungan antara penggunaan listrik dengan temperatur maka data temperatur yang akan diproses juga harus menyesuaikan dengan data penggunaan listrik. Dikarenakan data penggunaan listrik tersedia untuk per bulan, maka data temperatur juga harus diubah menjadi data bulanan.



Gambar 4. Gambar fp tree dari data record yang sudah terurut.

Tabel 1.
Data Record

No.	Record	No.	Record
1	TR PTD PU JD LT	13	TS PTJ PID JD LR
2	TR PTD PID JJ LR	14	TS PTD PID JJ LR
3	TR PTJ PID JD LR	15	TS PTD PU JD LT
4	TS PTD PU JD LT	16	TR PTD PID JJ LR
5	TS PTD PID JJ LR	17	TR PTJ PID JD LT
6	TS PTJ PID JD LR	18	TR PTD PU JD LT
7	TT PTD PID JJ LR	19	TS PTJ PID JD LR
8	TT PTJ PID JD LR	20	TS PTD PU JD LT
9	TT PTD PU JD LT	21	TS PTD PID JJ LR
10	TR PTJ PID JD LR	22	TS PTD PU JD LT
11	TR PTD PID JJ LR	23	TS PTD PID JJ LR
12	TR PTD PU JD LT	24	TS PTJ PID JD LR

Keterangan :

1. Temperatur rendah (TR)
2. Temperatur sedang (TS)
3. Temperatur tinggi (TT)
4. Pertanian jauh (PTJ)
5. Pertanian dekat (PTD)
6. Perikanan jauh (PUJ)
7. Perikanan dekat (PID)
8. Jalan jauh (JJ)
9. Jalan dekat (JD)
10. Listrik rendah (LR)
11. Listrik tinggi (LT)

Sama seperti data penggunaan listrik, akan dikenakan fungsi rata-rata untuk mengubah data temperatur harian ini menjadi data bulanan.

$$\mu = \frac{\sum \text{suhrata} - \text{rata}}{\text{jumlah hari}} \quad (6)$$

Untuk tanggal yang kosong dikarenakan tidak ada data atau datanya tidak terukur pada tanggal tersebut. Untuk mengatasi hal itu, maka tanggal yang kosong tersebut dihilangkan sehingga tidak terhitung ke jumlah hari melainkan hanya tanggal yang tersedia saja.

Setelah memperoleh temperatur rata-rata untuk semua bulan maka selanjutnya akan dibagi menjadi 3 kelas, yaitu rendah, sedang, dan tinggi.

3. Data jarak fitur geografi

Terdapat 3 fitur geografi yang akan dipakai dalam penelitian ini, yaitu : daerah pertanian, perikanan, dan jalan. Selanjutnya masing-masing fitur geografi tersebut dibagi menjadi 2 kategori yaitu dekat dan jauh. Penentuan kategori tersebut berdasarkan yang diperoleh dari GIS.

Tabel 2.
Daftar frequent pattern

Item	Frequent pattern generated
PUJ	(PTD, PUJ : 8), (JD, PUJ : 8), (LT, PUJ : 8), (PTD, JD, PUJ : 8), (PTD, LT, PUJ : 8), (JD, LT, PUJ : 8), (PTD, JD, LT, PUJ : 8)
JJ	(PTD, JJ : 8), (PID, JJ : 8), (LR, JJ : 8), (PTD, PID, JJ : 8), (PTD, LR, JJ : 8), (PID, LR, JJ : 8), (PTD, PID, LR, JJ : 8)
PTJ	(PID, PTJ : 8), (JD, PTJ : 8), (PID, JD, PTJ : 8)
LT	(PTD, LT : 8), (JD, LT : 8), (PTD, JD, LT : 8)
TR	0
TS	(PTD, TS : 8), (JD, TS : 8), (PID, TS : 8), (LR, TS : 8), (PID, LR, TS : 8)
LR	(PTD, LR : 8), (PID, LR : 15), (PTD, PID, LR : 8)
JD	(PTD, JD : 8), (PID, JD : 8)
PID	(PTD, PID : 8)
PTD	0

Tabel 3.
Daftar association rule yang diperoleh

No.	Rule	Support	Confidence	Lift
1	PID => LR	62,5%	93,75%	1.5
2	PTD => LR	33,33%	50%	0.8
3	PTD dan PID => LR	33,33%	100%	1.6
4	TS => LR	33,33%	66,67%	1.0667
5	PID dan TS => LR	33,33%	100%	1.6
6	JD => LT	37,5%	56,25%	1.5
7	PTD => LT	33,33%	50%	1.3333
8	PTD dan JD => LT	33,33%	100%	2.6665
9	JJ => LR	33,33%	100%	1.6
10	PTD dan JJ => LR	33,33%	100%	1.6
11	PTD dan PID dan JJ => LR	33,33%	100%	1.6
12	PID dan JJ => LR	33,33%	100%	1.6
13	PUJ => LT	33,33%	100%	2.6665
14	PTD dan PUJ => LT	33,33%	100%	2.6665
15	PTD dan JD dan PUJ => LT	33,33%	100%	2.6665
16	JD dan PUJ => LT	33,33%	100%	2.6665

$$\text{med} = \frac{\min(x_1, x_2, x_3) + \max(x_1, x_2, x_3)}{2} \quad (7)$$

dimana :

- x1 : jarak dengan rayon krian
x2 : jarak dengan rayon porong
x3 : jarak dengan rayon sidoarjo kota

Karena pembagian untuk masing-masing fitur geografi sama, maka diberlakukan cara yang sama. Untuk kategori dekat ($\leq \text{med}$) dan jauh ($> \text{med}$).

A. Proses Asosiasi

Pada bagian ini akan dijabarkan tentang alur berjalannya sistem. Seperti yang terlihat pada 1, sistem akan dimulai dengan mencari nilai *support* dari masing-masing item yang dilakukan dengan cara memindai data dari basis data. Nilai *support* ditentukan dari banyaknya kemunculan setiap item di keseluruhan *record* yang ada.

Setelah jumlah kemunculan atau nilai *support* dari masing-masing item didapatkan, selanjutnya akan dilakukan *sorting* atau pengurutan secara *descending*. Jadi item yang memiliki

nilai *support* tertinggi berada di urutan yang paling atas, begitu seterusnya sampai item yang memiliki nilai *support* terendah berada di urutan terbawah.

Selanjutnya setelah didapatkan urutan item berdasarkan nilai *support*nya secara keseluruhan akan dilakukan pengurutan lagi, tetapi yang diurutkan kali ini adalah item-item yang ada di setiap *record*. Urutannya sesuai dengan urutan item yang diperoleh sebelumnya. Hal ini dikenakan pada setiap *record* yang tersedia.

Langkah selanjutnya adalah membentuk *FP Tree*. Caranya yaitu dengan mengubah masing-masing *record* menjadi sebuah *path* yang nantinya akan membentuk *tree* setelah semua *path* digabungkan.

Proses asosiasi bisa dilakukan setelah *FP Tree* sudah terbentuk. Untuk selanjutnya akan dilakukan ekstrak *frequent pattern* berdasarkan *FP Tree* tersebut. Ada beberapa langkah untuk mengekstrak *frequent pattern* dari *Tree*. Langkah yang pertama adalah mendapatkan *conditional pattern base* yang caranya dengan melihat pola-pola apa saja yang sudah terbentuk dengan memilih suatu item untuk dijadikan sebagai acuan. Item acuan tersebut dipilih mulai dari item yang menempati posisi terbawah atau anak terakhir. Begitu seterusnya sampai atribut yang mendekati *root* atau *null*.

Langkah kedua adalah membentuk *conditional fp tree*. *Conditional fp tree* (CFPT) ini dibentuk berdasarkan CPB yang diperoleh sebelumnya, bukan berdasarkan basis data awal lagi.

Langkah berikutnya kita bisa mendapatkan *frequent pattern* yang sesuai dengan CFPT tersebut. Baru setelah memperoleh *frequent pattern* kita bisa menemukan *rule* yang mungkin terbentuk dengan cara mengkombinasikan seluruh item-item yang terdapat di masing-masing *frequent pattern* tersebut menjadi bentuk $A \Rightarrow B$ dengan A dan B terdiri dari satu atau lebih item dan item yang termasuk di A tidak termasuk di B atau dengan kata lain A dan B tidak memiliki irisan.

IV. HASIL PENELITIAN

Seperti pada Gambar 2 dengan transparency yang sesuai kita bisa membuat poligon untuk membatasi daerah Sidoarjo. Lalu membagi daerah Sidoarjo menjadi 3 bagian dengan menggunakan line sesuai dengan pembagian wilayah rayon dari PLN.

Fitur-fitur geografi yang dibutuhkan juga ikut dibuat sehingga nantinya akan terbentuk beberapa layer. Setelah itu menentukan titik pusat dari masing-masing rayon dan fitur geografi tersebut. Penentuan jarak diukur dari titik pusat rayon ke titik pusat fitur geografi seperti pada Gambar 3.

Setelah jarak masing-masing rayon dengan masing-masing fitur geografi didapatkan, maka seperti yang sudah dijelaskan sebelumnya akan dihitung mediannya. Nilai kurang dari atau sama dengan median dikategorikan dekat dan lebih dari median dikategorikan jauh.

Dalam sub bab ini dijelaskan salah satu hasil dari penelitian yang sudah dilakukan dan mewakili semua percobaan. Dengan menggunakan data 8 bulan dan minimal *support* 30% (7,2).

Setelah menghitung nilai *support* masing-masing item dan diurutkan lalu item yang ada di setiap *record* juga harus

diurutkan sesuai urutan per item sebelumnya. Sehingga *fp tree* bisa dibentuk dari gabungan semua *path* yang mana setiap *path* mewakili satu *record*.

Untuk selanjutnya akan dilakukan ekstrak *frequent pattern* berdasarkan *fp tree* tersebut. Ada beberapa langkah untuk mengekstrak *frequent pattern* dari *tree*. Langkah yang pertama adalah mendapatkan *conditional pattern base* yang caranya dengan melihat pola-pola apa saja yang sudah terbentuk dengan memilih suatu item untuk dijadikan sebagai acuan. Item acuan tersebut dipilih mulai dari item yang menempati posisi terbawah atau anak terakhir. Begitu seterusnya sampai item yang mendekati akar atau *null* (Gambar 4).

Setelah itu hitung jumlah kemunculan setiap item yang ada pada *conditional pattern base* masing-masing item acuan. Dan item yang memiliki *support* dibawah nilai *support* akan dihilangkan. Sehingga kita bisa mendapatkan *frequent pattern* nya. (Tabel 2)

Merujuk pada tujuan penelitian ini yaitu untuk mengetahui pola hubungan antara penggunaan listrik dengan temperatur dan fitur geografi, maka FPG yang akan diamati dan dianalisis yaitu FPG yang mempunyai atribut LR atau LT.

Setelah menemukan *frequent pattern* pada Tabel 2, selanjutnya akan digenerate *rule-rule* yang bisa dibentuk berdasarkan *frequent pattern* tersebut (Tabel 3).

V. ANALISA DAN KESIMPULAN

Maka dari percobaan di atas didapat 5 *association rule* yang berhubungan dengan fitur geografi dan selalu mempunyai nilai *lift ratio* tertinggi, serta 2 *association rule* yang berhubungan dengan temperatur, yaitu:

1. Jika jarak dengan pertanian dan jalan besar dekat maka konsumsi listrik tinggi.
2. Jika jarak dengan perikanan jauh maka konsumsi listrik tinggi.
3. Jika jarak dengan pertanian dekat dan dengan perikanan jauh maka konsumsi listrik tinggi.
4. Jika jarak dengan pertanian dan jalan besar dekat dan dengan perikanan jauh maka konsumsi listrik tinggi.
5. Jika jarak dengan jalan besar dekat dan dengan perikanan jauh maka konsumsi listrik tinggi.
6. Jika temperatur sedang maka konsumsi listrik rendah.
7. Jika jarak dengan perikanan dekat dan temperatur sedang maka konsumsi listrik rendah.

Dari *rule* yang terbentuk terdapat 3 macam pengertian. Dilihat dari nilai *lift rationya*, jika nilainya < 1 maka *rule* tersebut berkorelasi negatif atau dengan kata lain kemunculan salah satu item A atau B mempengaruhi hal yang sebaliknya pada kemunculan item lainnya. Jika nilai *liftnya* $= 1$ maka A dan B tidak ada korelasi atau bisa dikatakan *independent*. Dan jika bernilai > 1 maka *rule* ini dikatakan berkorelasi positif yang berarti kemunculan salah satu A atau B berhubungan dengan kemunculan yang lainnya.

DAFTAR PUSTAKA

- [1] S. Jovanovic, S. Savic, M. Bojic, Z. Djordjevic, and D. Nikolic, "The impact of the mean daily air temperature change on

- electricity consumption,” *Energy*, vol. 88, pp. 604–609, 2015.
- [2] A. Kipping and E. Tromborg, “Hourly electricity consumption in Norwegian households-Assessing the impacts of different heating systems,” *Energy*, vol. 93, pp. 655–671, 2015.
- [3] M. . Wijaya and T. Tezuka, “A comparative study of households’ electricity consumption characteristics in Indonesia: A techno-socioeconomic analysis,” *Energy Sustain. Dev.*, vol. 17, pp. 596–604, 2013.
- [4] R. . Rathod and R. . Garg, “Regional electricity consumption analysis for consumers using data mining techniques and consumer reading data,” *Electr. Power Energy Syst.*, vol. 78, pp. 368–374, 2016.
- [5] Kusriani and E. . Luthfi, *Algoritma Data Mining*. Yogyakarta: Andi, 2009.